2022 年北京冬奥会“全景看冬奥”VR 系统技术详解

汪念鸿 谢 胜 张佳 王 宇 唐俊权
（央视国际网络有限公司，北京 100142）

摘 要：央视网在 2022 年北京冬奥会期间推出的“全景看冬奥”节目，采用了自主研发的全套 VR 系统。文章详细阐述了适用于北京冬奥会 180° 鱼眼视频源的等角金字塔（EAP）映射模式、视频生产方案、基于直播中心的安全可靠分发调度系统、播放器，页面交互和渲染技术，以及 CDRM 版权保护系统。

关键词：北京冬奥会；VR 系统；EAP；视频生产与直播调度　　**中图分类号：**TP2　　**文献标识码：**A
文章编号：1671-0134（2022）04-021-06　　DOI：10.19483/j.cnki.11-4653/n.2022.04.004
本文著录格式：汪念鸿，谢胜，张佳，王宇，唐俊权.2022 年北京冬奥会“全景看冬奥”VR 系统技术详解 [J]. 中国传媒科技，2022（04）：21-26.

1. VR 系统概述

传统的 VR 系统由视频源、视频投影编码、视频传输和视频终端显示 4 个部分组成。结合具体业务需求，央视网自主研发的 VR 系统在每个部分都加入了新的技术（如图 1 所示）。针对北京冬奥会的 4K 180° Fisheye 视频源，央视网研发了专用的 EAP 映射模式，在保证高清晰度的同时，可节约一半以上的传输带宽。此外，央视网设计了高效的视频生产转码方案，提高视频生产速度，确保实效性；在分布式直播中心上设计调度矩阵，确保当播放的主码流出现问题时，及时调用备用流无缝切换，保障 VR 直播安全；为了给用户带来极致的 VR 视频观看体验，央视网在播放器中增加了包括大数据、广告和 VDN 等在内的 API，页面交互功能、空间渲染技术等。最后，为了达到本届冬奥会对版权保护的高标准要求，央视网采用了自主研发的 CDRM 对 VR 内容进行版权保护。



图 1　VR 直点播系统图

2.EAP 映射模式

2.1　VR 映射模式概述

VR 映射模式，用于编码之前将 360° 的球面图像映射成为二维平面图像。常用的 VR 映射模式有等距圆柱映射（Equi-Rectangular Projection，ERP）、立方体映射（Cube Map Projection，CMP）、等角立方体映射（Equi-Angular Cubemaps Projection，EAC）、金字塔棱台投影（Truncated Square Pyramid Projection，TSP）等，它们各有其优缺点。例如，目前应用最为广泛的 ERP，具有简单、计算快、转换后的二维平面比较直观，易于观看的优点，但其缺点是采样不均、在两级冗余大、不利于存储和传输。后续研究者提出的 CMP，虽然很好的改善了 ERP 的问题，但其在每一面的边缘仍然具有采样不均的缺点。2016 年，Facebook 提出的 TSP 采用截断金字塔的六面体来展示 360° 视频，在需要突出 VR 正面、弱化其它五面的任务上展现出比较好的效果，该方式可以很大程度地降低视频所需的存储空间，同时保证了展示面具有很高的清晰度。2017 年，谷歌提出了 EAC，进一步改善了 CMP 的问题，但却相对增大了计算消耗。

2.2　EAP 映射模式

2022 年北京冬奥会在 VR 直、点播中使用的是 4K 超高清 180° Fisheye 视频源（如图 2 所示），重点区域是整个 360° 视角范围内的前方区域，两侧及后方区域均为黑景，如果采用 CMP 或者 EAC 进行投影，即使使用紧凑布局，也只能将侧面、后面与正前方作为均等的画面来分割，作为主要区域的正前方其实只分配到了整个画面六分之一的分辨率。而在同等传输码率的情况下，使用 TSP，则能让正前方分配到整个画面二分之一的分



图 2　4K 超高清 180°　Fisheye 视频源图

辨率，极大地增强了主要关注区域的清晰度。同时，考虑到进一步拓展前面区域的视角范围和增强清晰度，央视网在 TSP 的基础上，结合了 EAC 采样均匀的优势，研发出等角金字塔（Equi-Angular Truncated Square Pyramid projection，EAP）映射模式，即，直接在截头方形金字塔（TSP）二分之一的正面视角及相邻的左右两侧进行等角度（Equi-Angular）形变。以此让同样大小的区域内能容纳更多的画面，而且，减少了通常逻辑下先转为立方体模型再进行投影的计算量。

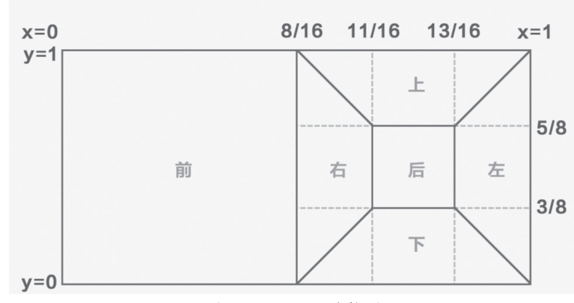


图 3　TSP 六面结构图

如图 3 所示为 TSP 六面结构的划分。在此基础上，EAP 针对六个面的二维平面到三维平面的映射采用了自主推导过程，其中前面和右面的推导公式如下所示。

面索引	二维到三维映射公式
	x=(i+0.5)/W i∈[0,W)，W，图像宽度 y=(j+0.5)/H j∈[0,H)，H，图像高度
前面	x'=tan(π/4.0*(x*4.0-1.0)) y'=tan(π/4.0*(y*2.0-1.0)) z'=1.0 x∈[0,0.5) y∈[0,1.0)
右面	x'=1.0 y'=tan(π/4.0*(2.0*(y-2.0*x+1.0)/(3.0-4.0*x)-1.0)) z'=tan(π/4.0*((8.0*(0.5-x)/(3.0-12.0*(x-0.5)))+1.0)) x∈[0.5,0.6875) y∈[0.0,0.375)∩[2.0*(x-0.5),0.375)∪[0.375,0.625)∪[0.625,1.0)∩[0.625,2.0*(1.0-x)]

i，j 表示在二维图像中像素位置，x，y 表示在归一化后的平面坐标系中的横坐标与纵坐标，取值范围是 [0，1]，x'，y'，z' 表示通过推导公式计算得到的三维坐标系中的坐标值。EAP 正面和侧面的映射如图 4 所示，EAP 正面增加了等角度变换，使得正面的像素分配更均匀，EAP 侧面采用了自主推导算法加等角变换。

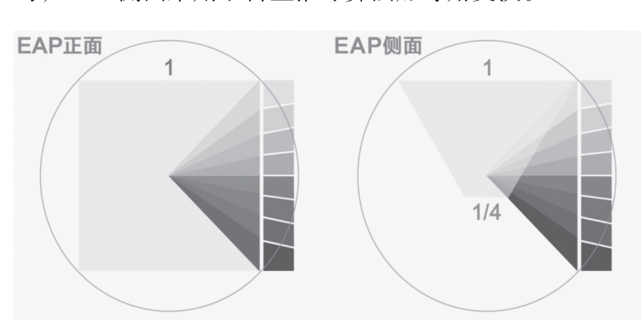


图 4　EAP 正面与侧面映射示意图

图 5 所示，是 EAP 格式的 VR 视频图。央视网采用 EAP 在对 VR 视频进行转码的同时还加入了自主研发的极速高清编码优化技术，使得输出的 2K（2880x1440）视频源与 4K 的清晰度基本一致，但节约带宽一半以上。



图 5　EAP 格式视频截图

3. VR 点播视频生产调度系统

3.1　VR 视频生产调度方案

为了完成北京冬奥会 VR 点播视频生产需求，央视网北京即刻系统与无锡视频生产系统协同工作，经无锡内容团队审核后的 4K 超高清 180° 的 Fisheye 清流视频通过专线传输到北京，即刻系统开放接口接收视频基本信息和编目信息，最终进入转码能力池进行 VR 视频转码（如图 6 所示）。为了提升 VR 视频的转码效率，新增 4 台高性能转码服务器，同时采用 SIMD、预查表优化方式和并行算法对 EAP 投影计算过程进行了优化，将 VR 视频的生产效率提升了 7–8 倍，完全达到了本届冬奥会大量 VR 视频的制作时效性要求。

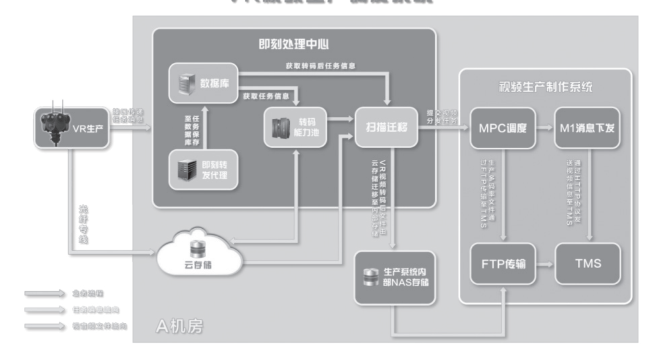


图 6　VR 视频生产调度系统图

3.2　即刻视频转码处理系统概述

即刻系统是央视网自主研发的视频标准化处理平台，支持各种格式的视频文件上传进行并发转码，转换成生产系统能够处理的视频格式。该系统采用了最新技术架构，严格按标准规范设计，具有软件高扩展性，系统易用、易维护，高稳定性等优点。即刻系统架构图如图 7 所示，下面主要从 4 个部分进行介绍。

3.2.1　Web 上传

即刻系统采用最新的前后端分离技术，后端采用Java架构实现高并发视频上传、接口采用Restful接口规范，实现包含VR在内的各种编码格式和封装视频上传，并保存视频信息和转码策略至后台数据库。支持批量视频文件和大并发文件上传，上传文件保存至私有云存储。为保障北京冬奥会4K超高清180° Fisheye视频正常制作，开发新接口接收视频基本信息和编目信息，无锡视频生产系统通过专线将VR视频传输至私有云存储，并由转码服务集群进行转码处理。

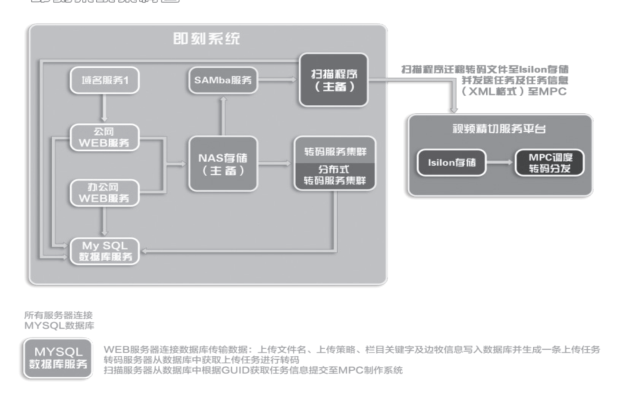


图7　即刻系统架构图

3.2.2　视频转码

VR视频转码采用自主研发的VR编码器，采用SIMD、预查表等优化方式和并行算法对贴图计算进行了优化，转码效率得到大幅提升。对于每一条视频，设置不同的转码策略，由用户选择转码策略，连同视频的编码信息保存至数据库，同时部署N台转码服务构建即刻转码能力池进行视频转码并行处理，对视频任务设置任务优先级保证重要视频第一时间进行处理。每台转码服务都可从数据库获取一条视频任务进行处理，同时对数据库采用锁竞争保护机制，如果任务被一台转码服务执行，其他转码服务不再有权利执行。对长视频编码，为了提高转码效率，采用Hadoop的Map/Reduce机制，将编码任务拆分到各机器进行高效地处理转码。为避免音频数据切片时产生爆音，采用对音视频分离处理的方式，只对视频进行拆分、编码和合并，把处理完成的视频按照先后顺序进行合并，然后再与音频进行合并，这样在分片的接缝处不会出现爆音，完美解决了大文件视频转码速度慢的问题，并且可以根据业务需要扩充集群规模，不断提升转码效率。

3.2.3　生产系统调度

由即刻转码能力池处理过后的视频，通过扫描传输程序将转码后的视频文件迁移至生产系统内部存储，同时发送相应的任务制作消息至MPC调度消息队列，MPC从消息队列获取消息，对视频文件进行多码率转码和必要的包装，然后将各码率的最终的转码文件发送给下游TMS。

3.2.4　安全保障

所有服务器安装Zabbix agent，保证服务器的安全运行，利用其灵活的通知机制及时发现服务器状态的监视，出现问题保证了第一时间进行排查和解决。另外对各系统密码做到最高级别设置并定期修改，对网络安全风险漏洞及时进行修复，北京冬奥会期间做到了系统稳定、安全高效、全天不停歇运行，制作的VR视频清晰流畅。

4. VR直播调度系统

为了保证直播流的高可用性，避免单点故障引起的大范围停播事件，央视网基于最新建设的分布式直播中心（如图8所示），引入主、备流接入及无缝切换技术，建立了VR直播调度系统（如图9所示）。在该系统中，央视网使用不同厂商的编码器生成主、备两路或者多路直播流，通过Http-flv或Rtmp协议接入到直播中心的调度矩阵，调度矩阵与源站之间使用http-vflv协议连接。当调度矩阵接入的单路流发生链路异常、音视频时间戳不同步、视频时间戳异常等故障时，会自动在一个GOP时间内切换到备用流，并且发送新流的编码参数给边缘源站，边缘源站会启动新的切片逻辑，生成新的ts切片，通过hls协议送达用户终端。在切换过程中，调度矩阵始终与源站连接，用户不会出现卡顿、闪断等现象，极大地提升了用户体验。运营策略上，对视频流采用一对一、一对多的调度，实现同一路流对不同客户的分销策略；同时，采用子矩阵划分来实现业务逻辑清晰划分。这些巧妙的设计，大幅减少了系统开销，且支持定时任务，降低了运维成本。

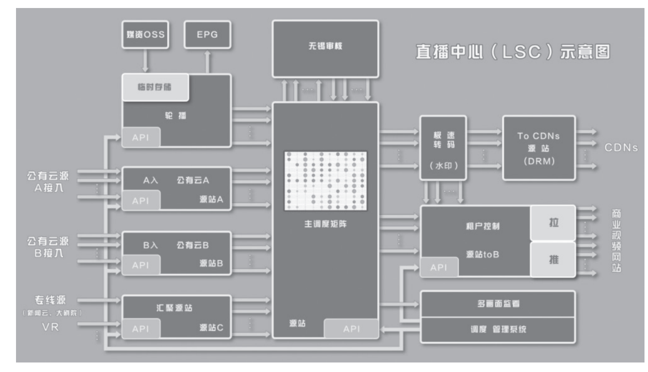


图8　直播中心示意图

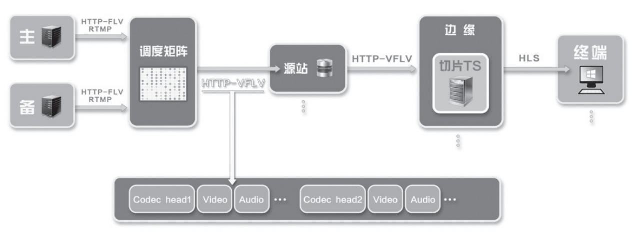


图9　VR直播调度系统

5. VR播放器

5.1　API

央视网自主研发的H5-VR播放器拥有功能完备的API，可以支持多种VR视频格式、多终端、多浏览器以及多种视频信息获取方式。

播放器可以适配 ERP、CMP、EAC、EAP 等多种格式的 VR 投影视频。页面集成时，可通过特定参数，来确定播放哪种格式的 VR 视频，满足了多样化的 VR 视频流播放需求。

5.1.1 支持多终端、多浏览器

播放器能支持在 PC 端的 Windows、Mac、Linux 和移动端的 Android，华为的鸿蒙和苹果的 iOS 这些被广泛应用的系统。PC 端的浏览器支持 Chrome、Safari、Firefox、Edge 等主流浏览器，移动端浏览器支持系统原生、微信的 X5 内核以及其他国内广泛使用的浏览器。央视网还一一做了专项适配，使得我们的 VR 播放器在国内适应和覆盖能力得到了大幅提高。为提高通用性，降低维护成本，央视网的 VR 播放器可自动识别是 PC 端还是移动端，并给出相应的呈现。相同的参数，既支持直播的频道 ID，也支持点播的 Guid，播放器识别后，请求对应的直播视频分发网络 VDN 和点播 VDN，从而获取对应的直播流和点播流地址。

5.1.2 支持多种视频信息获取方式

视频流地址、标题、播放器预播图片等视频信息，既可以通过参数直接传递，又能从 VDN 间接获取，满足了多样化的业务需求。

此外，播放器接受直点播视频分发网络（Video Delivery Network，VDN）和广告调度系统控制，集成了央视网大数据统计接口，对播放数据和播放质量进行了统计、收集和上报，方便对视频质量进行实时监控和调整。播放器还采用了统一的视频播放内核，使得页面集成时更加方便，只需要通过简单的参数设置，就能使移动端用像 PC 端一样的窗口播放，不仅能横向铺满屏幕播，还能横向、竖向同时铺满屏幕播。

5.2 页面交互

本届冬奥会引入的是 4K 源，为了给用户带来更好的观看体验，央视网为 VR 播放器新增了虚拟手柄和手势操作，PC 端可以通过鼠标拖动、滚轮缩放操作，移动端通过多种手指触摸手势，便捷地进行画面的方向调整以及虚拟镜头的推拉操作，还可以任意放大指定区域进行播放。画面放大时还能自动进行清晰度调整，使用户能看到画面细节。除这一常规的控制操作方式以外，央视网还创新加入虚拟摇杆设计，单指完成推拉摇移的操作，在 PC 端也可以用鼠标拖拉虚拟手柄进行类似操作，尤其是在平滑匀速移动观看上比拖动或滑动操作更具优势。以上技术使得 4K 超高清在移动小屏上得到了更好展现。另外，播放器还增加了迷你雷达指示视角的大小和方向，为 VR 的控制操作带来更为直观的标识，给一进 VR 就头晕的用户，提供了更好的交互体验。

5.3 空间渲染显示

VR 视频 EAP 格式的优点是占用资源和带宽比其他格式都少很多，缺点是单流情况下只有一个视区清晰，其他面都比较模糊。为解决此问题，央视网设计了可根据用户观看视角进行清晰度切换的方法，保证了用户观看的视角为高清晰画面，节约传输成本的同时最大程度提升了用户体验。切换清晰度方法可通过分流或者分层两种方式实现。

服务器端将 4K 超高清 360° VR 源视频通过 EAP 编码器编出 4~6 路 1080P 或者 2K 不同平面的视频流（如图 10 所示），用户在客户端通过拖动或陀螺仪选择观看的视区，当用户在某个视角停留时间超过设定时间（一般 1~3 秒左右），播放器就会根据此角度向服务器请求对应的视频流并加载播放，完成清晰度切换（如图 11 所示）。此方法有以下特性：整个视频流切换，不会产生角度偏差和变形等问题；如果新视角视频流加载失败，播放器会继续播放前一次已经切换成功的视频流；视频流切换完成时间跟单个 ts 的播放时长及网络速度有关，单 ts 播放时长越短，切换越快，这些时间参数设定都可根据实际情况进行调整。

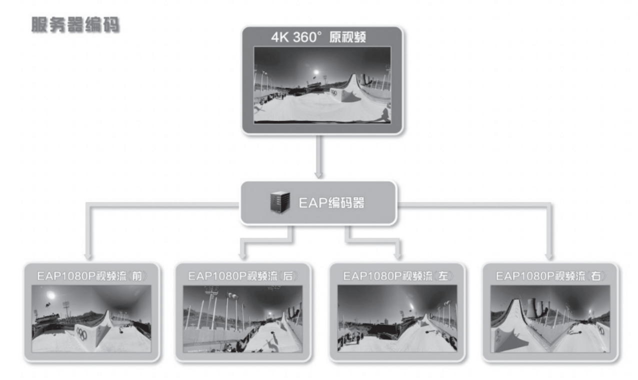


图 10 EAP 编码器分流示意图

清晰度切换对比





图 11 清晰度切换对比示意图

5.3.2 通过分层方式实现清晰度切换

首先，把 4K 超高清源视频编码成一个普通清晰度视频作为一个基础层，然后用它们之间的残差层再编译出不同角度的视频流。其次，客户端先拉取一个基础层的视频流，再根据当前用户观看视角拉取对应的残差层视频流，最后将两个视频合成，以此得到此观看区域的高清视频。

6. 版权保护系统

本届冬奥会对版权保护要求异常严格，包含了对播放终端、地域和播放时间三个维度的版权保护限制。央视网使用 ChinaDRM 标准与自主 CDRM（CNTV Digital Right Management）对视频流版权进行保护，

并实现两种模式的动态切换，以加固版权的保护。在播放终端，央视网研发了 WASM 扁平混淆编译器对解密和版权保护模块进行混淆编译加固。此外，版权管理人员还能对指定视频的授权播放地区和时间段进行版权配置，然后播放器根据 VDN 返回的版权字段对版权信息进行识别，确定当前地区和当前时间段是否有权播放此 VR 视频。

6.1 版权保护技术概述

先对比介绍 3 种版权保护技术的优缺点，包括防盗链、传统加密模式、数字版权管理系统（Digital Rights Management，DRM）。在此基础上，进一步介绍 CDRM 系统。

6.1.1 防盗链

央视网使用的第一代防盗链方案是给 m3u8 和 ts 文件的 url 动态生成一个 token，使得 url 会随着时间、IP 地址、用户 ID、内容 ID 等很多参数而动态变化，导致其他人无法使用这个 url 进行盗链。但是这种方式只能防止其他网站直接使用你的 url 来观看或者一般用户的下载。

6.1.2 传统加密模式

现在比较主流的加密方式是传统加密模式，给 ts 文件加密后，m3u8 通过 tag：#EXT-X-KEY 提供一个 url 就可以获取加密 key，然后对 ts 片段解密进行播放。这种加密模式的缺点是任何人都可以获取到 m3u8 内的 tag：#EXT-X-KEY 提供的 url，其基本原理就是使用第一种方案来加密了 #EXT-X-KEY 所指定的 url，其缺点不言而喻。

6.1.3 DRM 系统

DRM 系统，用户系统为用户或者设备颁发证书，认证系统先对用户进行认证；认证成功后，认证系统将内容密钥通过用户证书公钥加密送往客户端，高强度的非对称加密方式提高了破解密钥的难度；终端使用私钥对密文解密后获取内容密钥；通过以上两步最终完成 HLS 协议中指定的 ts 文件解密。

6.2 CDRM 版权保护系统

基于 DRM 系统，央视网研发的 CDRM 系统工作流程如图 12 所示，分为加密与解密两个部分。

6.2.1 传递内容密钥

加密系统向认证系统发起认证请求，认证系统对加密系统证书链进行认证，认证通过后使用加密系统证书公钥将内容密钥加密，发送给加密系统，最后加密系统使用私钥将密文解密。

6.2.2 使用内容密钥加密内容

加密系统对 ts 分片的音视频独立加密，对 H264/H265 的 NALU 加密，加密方式与模式可选，支持 DES、AES、MS1 等。可扩展的加密协议接口支持随意的组合加密方式与模式，在直播过程中可以做到实时调整。要得到可以播放的视频文件，还需要对加密逻辑反向分析，极大地增加了破解难度。通过以上两步最终完成 HLS 协议中指定的 ts 分片加密。

6.2.3 获取内容密钥

终端用户登录到用户系统，用户系统验证用户的合法性，然后为用户、设备颁发证书。认证系统对用户证书链先进行认证，成功后将内容密钥通过用户证书公钥加密送往客户端，客户端收到密文后使用私钥解密，最终获取内容密钥。

6.2.4 解密视频

终端通过解密系统发出请求，并带上证书信息。认证系统会对终端用户的证书发起多种鉴定，比如用户权限、证书真伪、过期时间等。如果通过认证，终端将会获取到对应内容密钥。这个内容密钥是用终端的公钥加密的，只能使用对应的私钥才能加密。最终，从源站获取的加密 ts 分片在解密系统中，按照对应的解密算法，使用内容密钥解密视频。

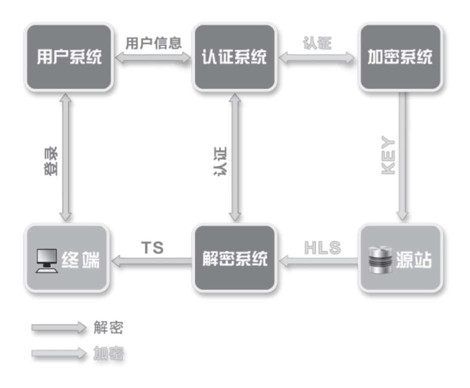


图 12 CDRM 系统

6.3 播放器终端版权保护

加密的视频流到达客户端后，需要被实时解密和解码来呈现给用户，这不但要求解密解码的速度得快，还要求解密模块抗破解能力要特别强。因此央视网的 H5 播放器解密模块采用了 LLVM 技术，即把解密的代码放入 WASM 模块中，并且通过十六进制的方式传输到用户页面，这有效地解决了在用户页面出现解密代码容易被跟踪调试的问题。同时，在编译 WASM 模块时，没有使用通用的 es 编译器，而是采用了自研的 WASM 扁平混淆编译器对解密和版权保护模块进行混淆编译加固，进一步提升其抗破解能力。自研的 WASM 扁平混淆编译器包括以下两项混淆技术。

6.3.1 虚假控制流混淆

如图 13 所示，从虚假控制流混淆对比前后可以看出，

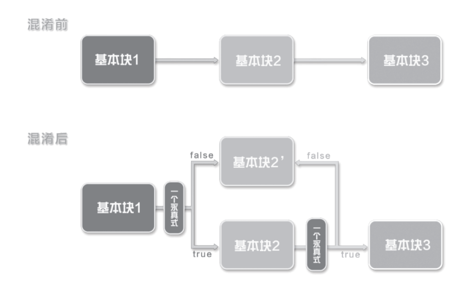


图 13 虚假控制流混淆对比图

相比于混淆前原程序中的单个代码块，混淆后的程序中加入了一个由不透明永真式组成的条件语句和一个永不达的混淆基本块 2 '。由于混淆器自身知道不透明永真式的值，所以可以保证被混淆程序执行原本的代码，保证了执行的正确性。而编译器和反编译器都无法对不透明永真式进行求值，只能保留此不透明永真式，达到干扰静态分析的目的。

6.3.2 控制流平坦化混淆

如图 14 所示，从控制流平坦化混淆对比前后可以看出，相比于混淆前原程序中的单个代码块，混淆后的程序中除基本块 1 外，其他的基本块都集中到了同一个层次上。不同基本块的调用顺序由主分发器决定，这样可以模糊基本块之间的前后关系，增加程序分析的难度。

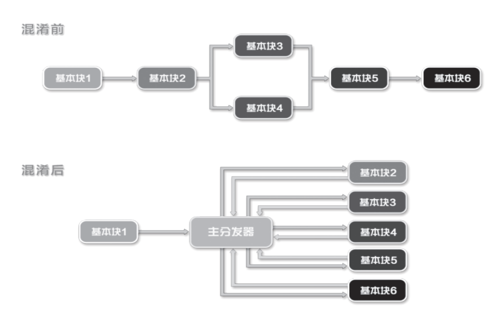


图 14 控制流平坦化混淆对比图

通过自研的混淆编译器生成的 WASM 文件，使得解密解码速度和抗破解能力达到最佳平衡，在大幅度提高了 WASM 反编译破解难度的同时，也保证了解密超高清节目的实时性。此外，央视网还对页面的 JavaScript 文件做了混淆处理，以此提升了进入分析 WASM 之前的破解难度。

结语

央视网采用一整套自主研发的 VR 系统为观众带来了一场别样的全景冬奥视听体验，为今后推出更多优质 VR 节目构建了技术先进、自主可控的技术支撑平台。为了实现更清晰，更节约成本，更快转码，更好交互体验等目标，央视网视频技术研发团队进行更深入的钻研，力求在 Simd 指令优化、GL 模型优化、编码标准多层优化、播放器虚拟场景等方面取得进一步突破。媒

作者简介：汪念鸿（1981-），男，湖北孝感，工程师，央视国际网络有限公司，研究方向：视频播放器及视频编解码；谢胜（1982-），男，四川德阳，工程师，央视国际网络有限公司，研究方向：视频传输与加密；张佳（1983-），男，河北衡水，工程师，研究方向：点播生产及视频编码；王宇（1979-），男，四川成都，工程师，央视国际网络有限公司，研究方向：H5 音视频播放；唐俊权（1985-），男，四川巴中，工程师，央视国际网络有限公司，研究方向：播放器及音视频传输。

（责任编辑：陈旭管）